



تمرین اول طراحی کامپایلر و  
زبان‌های برنامه‌نویسی



بهار ۱۴۰۴  
مهلت تحویل: ۱۴۰۳/۱۲/۲۳

دانشکده مهندسی برق و  
کامپیوتر - علوم کامپیوتر

طراحان تمرین: گلبو رشیدی، دریا انصاری‌پور، امیر فریدی

1. با فرض آنکه  $E = \{a^i b^j \mid i \neq j \text{ and } 2i \neq j\}$  باشد، نشان دهید  $E$  یک زبان مستقل از متن است. سپس برای رشته‌ی  $a^6 b^7$  درخت اشتقاق را رسم کنید.

2. با فرض آنکه الفبای زبان زیر به صورت  $\{0, 1\}$  باشد، یک گرامر غیرمبهم برای زبان زیر ارائه دهید.  
{w | in every prefix of w the number of 0's is at least the number of 1's}

3. طبق گرامر زیر تمامی مراحل اشتقاق رشته داده شده را بنویسید.  
رشته:

$-2 + a1 * (4-b)$

گرامر:

$S \rightarrow T + S \mid T - S \mid T$

$T \rightarrow P * T \mid P / T \mid P$

$P \rightarrow (S) \mid [S] \mid V \mid N$

$N \rightarrow M \mid -M$

$M \rightarrow [0-9]M \mid [0-9]$

$V \rightarrow [a-z]G$

$G \rightarrow [a-z]G \mid [0-9]G \mid \epsilon$

4. الفبای  $\Sigma_3$  را در نظر بگیرید. این الفبا تمام ستون‌های با اندازه‌ی 3 با درایه‌های 0 و 1 را شامل می‌شود.

$$\Sigma_3 = \{ (0\ 0\ 0), (0\ 0\ 1), (0\ 1\ 0), \dots, (1\ 1\ 1) \}$$

در این سوال  $(x, y, z)$ ، به معنی ماتریس ستونی زیر است:

x
y
z

زبان  $E$  را روی این الفبا به صورت زیر تعریف می‌کنیم:

$$E = \{ w \in \Sigma_3^* \mid \text{the bottom row of } w \text{ is the sum of the top two rows} \}$$

برای مثال:

$$(0\ 0\ 1) (1\ 0\ 0) (1\ 1\ 0) \in E, (0\ 0\ 1) (1\ 0\ 1) \notin E$$

نشان دهید زبان  $E$  یک زبان منظم است (کار کردن با وارون این زبان ممکن است راحت تر باشد). برای حل این سوال می‌توانید از نکته‌ی زیر استفاده کنید:

با فرض آنکه  $w = w_1 w_2 \dots w_n$  باشد، وارون رشته‌ی  $w$  به صورت  $w^R = w_n \dots w_2 w_1$  تعریف می‌شود.

حال زبانی مانند  $A$  را در نظر بگیرید. وارون این زبان را به صورت  $A^R = \{w^R \mid w \in A\}$  نشان می‌دهیم که  $A^R$  نیز یک زبان منظم است. کافیت با DFA این زبان، NFA-ای به این صورت بسازیم که حالت شروع آن، حالت‌های پذیرش DFA می‌باشد (اگر بیش از یک حالت پذیرش داشت، کافی است یک حالت جدید به عنوان حالت شروع تعریف کرده و آن را با  $\lambda$  به حالت‌های پذیرش وصل کنیم). و حالت شروع DFA اولیه، حالت پذیرش NFA می‌شود. حال برای تکمیل ماشین، تمام گذرهای ماشین اولیه را برعکس می‌کنیم. NFA-ای که به این صورت ساخته می‌شود، پذیرنده‌ی زبان  $A^R$  است.

5. گرامر زیر که زبان lambda calculus را تولید می‌کند را در نظر بگیرید.

$$M \rightarrow \text{var} \mid M M \mid \lambda \text{ var. } M \mid (M)$$

که در آن  $\lambda$  و "." و "(" و ")" پایانه هستند و var غیرپایانه‌ای است که identifier-ها را بدون ابهام تولید می‌کند.

الف) نشان دهید که این گرامر مبهم است و برای یک رشته، دو درخت اشتقاق متفاوت ارائه دهید.

ب) یک گرامر نامبهم ارائه دهید که دقیقاً رشته‌های گرامر داده شده را تولید کند.

6. فرض کنید در زبان C یک built-in data type به نام list برای نگهداری لیستی از اعداد صحیح تعریف شده است که ویژگی‌های زیر را دارد :

- این ساختار داده random access است یعنی به هر عضو آن میتوان با دستور زیر دسترسی داشت:

`L[index]`

- `L[index]` یک مقدار lvalue است، یعنی بطور مثال عمل زیر بر آن مجاز است :

`list L;`

`L[0] = 1;`

- مقداردهی اولیه آن در هنگام تعریف و به صورت زیر است :

`list L = [ 1, 2, 3, 4];`

`L = []; //an empty list`

- طول لیست را به صورت زیر میتوان بدست آورد :

`list L = [1, 2, 3, 4];`

`int a = L.length(); // this statement assigns length of L to a`

- مقدار index یک لیست میتواند بین 0 تا طول آن منهای ۱ باشد یعنی :

`list L = [1,2,3,4]`

`int a = L[ 4 ] /* this is a compile error, the number given as index must be greater than or equal to zero and less than L.length() */`

- میتوانیم در ابتدا و انتهای آن با syntax مقابل یک عدد insert کنیم:

`list L;`

`int a = 3;`

`list L1 = a + L; // this line inserts #3 to the beginning of the list`

`list L2 = L + a; // this line inserts #3 to the tail of the list`

توجه داشته باشید که نتیجه‌ی این عملگر لیستی است که فقط عنصر 3 به آن اضافه شده است. بنابراین میزان حافظه‌ی مصرفی آن برابر میزان حافظه‌ی مصرفی برای L به اضافه‌ی ۲ (برای افزودن عضو جدید) است.

(الف) به صورت declarative برنامه‌ای بنویسید که عضو اول و آخر یک لیست را swap کند.

(ب) به صورت imperative برنامه قسمت (الف) را بازنویسی کنید.

(ج) میزان حافظه‌ی مصرفی توسط این دو روش را بدست آورید و با یکدیگر مقایسه کنید.

7. برای هر یک از موارد زیر NFA یا DFA رسم کنید و عبارت منظم مربوط به هر یک را بنویسید.  
الف) تمام رشته‌هایی که عدد دودویی آن بر 6 بخشپذیر است.  
ب) تمام رشته‌هایی که زیر رشته‌ای به شکل 110 دارند اما زیر رشته 000 ندارند.  
ج) زبانی که رشته زمان، به صورت دوازده ساعته با دقت دقیقه (برای مثال رشته "7:07AM") را بپذیرد. در اینجا دقیقه باید به صورت عددی دو رقمی باشد و ساعت می‌تواند تک رقمی باشد.

8. رشته‌ی 11001001010101001101 را بر اساس زبان‌های توصیف شده در زیر توکن بندی کنید.  
الف) مجموعه اعداد دودویی که شامل زیر رشته 101 نباشند.  
ب) مجموعه اعداد دودویی بخش پذیر بر 4  
ج) رشته‌های دودویی که تعداد 0های آن زوج و تعداد 1های آن فرد باشد.  
در ابتدا فقط یک زبان را در نظر گرفته و در انتها همه‌ی این زبان‌ها را در نظر بگیرید. یعنی ابتدا فقط زبان قسمت‌های توصیف شده در بالا را به تنهایی در نظر بگیرید و سپس تمام زبان‌ها را به صورت همزمان در نظر بگیرید. در صورت نیاز برای رفع خطا از روش Panic Mode استفاده کنید.

#### توضیحات

- یک فایل به نام HW1-SID.pdf را آپلود کنید که SID شماره دانشجویی شما می‌باشد.
- در صورت تشخیص شباهت و تقلب میان حل تمرین شما و دیگران، نمره صفر به هر دو طرف تعلق خواهد گرفت.

موفق باشید!